Section 14.5: Multivariate Chain Rule:

Goal: Extend the chain / composition rule for derivatives from Calc I into Calc III.

In Calc I,

$\mathbb{R} \xrightarrow{f} \mathbb{R} \xrightarrow{g} \mathbb{R}$

$(g \circ f)(x) = g(f(x))$

In Calc III,

$\begin{pmatrix} \mathbb{R}^n \xrightarrow{f} \mathbb{R} \\ \mathbb{R}^k \xrightarrow{g} \mathbb{R} \end{pmatrix}$ how to compose?

Composition of Multivariate Functions:

Given $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$.

Then f has expression $f(x_1, x_2, \ldots, x_n)$.

Letting $g_i(t_1, t_2, \ldots t_n)$ for $1 \leq i \leq n$

We can define the composition of f w/ the g_i's

__via:__ $f(g_1(t_1, t_2, \ldots t_k), g_2(t_1, t_2, \ldots t_k), \ldots, g_n(t_1, \ldots, t_k))$

Ex: Suppose $f(x, y, z) = \cos(x+y) z^2 + 3$.

$x(s,t) = s+t$, $y(s+t) = st$, $z(s,t) = \cos(s)$

$f(x(s,t), y(s,t), z(s,t))$

$= f(s+t, st, \cos(s))$

$= \cos((s+t) + st)(\cos(s))^2 + 3)$ ▨.

Picture:

$$\begin{tikzcd}
\mathbb{R}^k \arrow[r, "g_1"] \arrow[dr, "g_2"] \arrow[ddr, "g_n"'] & \mathbb{R} \arrow[dr, "1^{st}\text{ coor.}"] & & \\
& \mathbb{R} \arrow[r, "2^{nd}\text{ coor.}"] & \mathbb{R}^n \arrow[r, "f"] & \mathbb{R} \\
& \mathbb{R} \arrow[ur, "n^{th}\text{ coor.}"'] & &
\end{tikzcd}$$

"shortcutting" the coords. yeilds:

$\mathbb{R}^k \xrightarrow{(g_i)} \mathbb{R}^n \xrightarrow{f} \mathbb{R}$

Now, we'll try to make the goal happen.

Setup: Let $f(x,y)$ and $x(t), y(t)$ be differentiable functions.

Def^n: A function $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ is differentiable at p when f is "well-approximated" by its tangent (hyper) plane near p.

(i.e. the error approximating f by its tangent plane near p goes to 0 as $\vec{x} \to \vec{p}$).

Now, given $f, x,$ and y as above with $p = (a,b)$
$f(x,y) = f(a,b) + (f_x(a,b) + \mathcal{E}_x(x,y))(x-a) + (f_y(a,b) + \mathcal{E}_y(x,y))(y-b)$
where $\mathcal{E}_x$ and $\mathcal{E}_y$ are error terms with $(\mathcal{E}_x, \mathcal{E}_y) \to (0,0)$ as $(x,y) \to (a,b)$.

$$\therefore f(x,y) - f(a,b) = (f_x(a,b))(x-a) + (f_y(a,b))(y-b) + \mathcal{E}_x(x-a) + \mathcal{E}_y(y-b)$$

Choose a time α where $(x(\alpha), y(\alpha)) = p = (a,b)$

Subsitute into the function to obtain:
$$f(x(t), y(t)) = f(x(\alpha), y(\alpha)) = f(x(\alpha), y(\alpha))(x(t) - x(\alpha)) + f_y(x(\alpha), y(\alpha))(y(t) - y(\alpha)) + \mathcal{E}_x(x(t) - x(\alpha)) + \mathcal{E}_y(y(t) - y(\alpha))$$

For each $t \neq \alpha$ we divide by $t - \alpha$ to obtain:
$$\frac{f(x(t), y(t)) - f(x(\alpha), y(\alpha))}{t-\alpha} = f(x(\alpha), y(\alpha))\left(\frac{x(t) - y(\alpha)}{t-\alpha}\right) + f_y(x(\alpha), y(\alpha))\left(\frac{y(t) - y(\alpha)}{t-\alpha}\right) + \mathcal{E}_x\left(\frac{x(t) - x(\alpha)}{t-\alpha}\right) + \mathcal{E}_y\left(\frac{y(t) - y(\alpha)}{t-\alpha}\right)$$

Limiting $t \to \alpha$ we obtain:

$$\frac{d}{dt}\left[f(x(t),y(t))\right]\Big|_p = \lim_{t\to\alpha} \frac{f(x(t),y(t)) - f(x(\alpha),y(\alpha))}{t-\alpha}$$

$$= f_x(x(\alpha),y(\alpha)) \lim_{t\to\alpha} \frac{x(t)-x(\alpha)}{x-\alpha} + f_y(x(\alpha),y(\alpha)) \lim_{t\to\alpha} \frac{y(t)-y(\alpha)}{t-\alpha}$$

$$+ \lim_{t\to\alpha} \varepsilon_x \cdot \lim_{t\to\alpha} \frac{x(t)-x(\alpha)}{t-\alpha} + \lim_{t\to\alpha} \varepsilon_y \cdot \lim_{t\to\alpha} \frac{y(t)-y(\alpha)}{t-\alpha}$$

$$= f_x(x(\alpha),y(\alpha))x'(\alpha) + f_y(x(\alpha),y(\alpha))y'(\alpha) + \cancel{\lim_{t\to\alpha}\underset{0}{\varepsilon_x}}\, x'(\alpha) + \cancel{\lim_{t\to\alpha}\underset{0}{\varepsilon_y}}\, y'(\alpha)$$

$$= f_x(x(\alpha),y(\alpha))x'(\alpha) + f_y(x(\alpha),y(\alpha))y'(\alpha)$$

Generalizing a little bit would yeild the following:

Prop: (Multivaritable Chain Rule): Let $f(x_1, x_2, \dots x_n)$ and $x_i(t_1, t_2, \dots t_k)$ be diff for $1 \le i \le n$. Then

$$\frac{\partial f}{\partial t_j} = \frac{\partial f}{\partial x_1} \cdot \frac{\partial x_1}{\partial t_j} + \frac{\partial f}{\partial x_2} \cdot \frac{\partial x_2}{\partial t_j} + \dots + \frac{\partial f}{\partial x_n} \cdot \frac{\partial x_n}{\partial t_j}$$

for all $1 \le j \le k$.

Comment: Definitely can't cancel ∂x_i's ... That would invalidate the formula!

Ex: Compute $\frac{\partial f}{\partial s}, \frac{\partial f}{\partial t}$ for $f(x,y) = e^x \sin(y)$, $x = st^2$, $y = s^2t$

Sol1 (w/o chain rule): First we compute composition

$f(x(s,t), y(s,t)) = f(st^2, s^2t) = \exp(st^2)\sin(s^2t)$

$\therefore \frac{\partial f}{\partial s} = \frac{\partial}{\partial s}[\exp(st^2)\sin(s^2t)] = \frac{\partial}{\partial s}[\exp(st^2)]\sin(s^2t) + \exp(st^2)\frac{\partial}{\partial s}[\sin(s^2t)] = t^2\exp(st^2)\sin(s^2t) + \exp(st^2)\cdot 2st\cos(s^2t)$

$\frac{\partial f}{\partial t} = \frac{\partial}{\partial t}[\exp(st^2)]\sin(s^2t) + \exp(st^2)\frac{\partial}{\partial t}[\sin(s^2t)] = 2ts\exp(st^2)\sin(s^2t) + \exp(st^2)\cdot s^2\cos(s^2t)$

Sol2 (w/ Chain Rule): To compute the desired partials:

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial s} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial s} \text{ and } \frac{\partial f}{\partial t} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial t} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial t}$$

Plugging into these equations

$$\frac{\partial f}{\partial x} = e^x \sin(y) = \exp(st^2)\sin(s^2 t)$$

$$\frac{\partial f}{\partial y} = e^x \cos(y) = \exp(st^2)\cos(s^2 t)$$

we got to reuse these ü

$$\frac{\partial x}{\partial s} = t^2 \qquad \frac{\partial y}{\partial s} = 2st$$

$$\frac{\partial x}{t} = 2st \qquad \frac{\partial y}{\partial t} = s^2$$

$$\frac{\partial f}{\partial s} = \exp(st^2)\sin(s^2 t) \cdot t^2 + \exp(st^2)\cos(s^2 t) \cdot 2st$$

$$\frac{\partial f}{\partial t} = \exp(st^2)\sin(s^2 t) \cdot 2st + \exp(st^2)\cos(s^2 t) \cdot s^2$$ ▨.

* Exercise: Let $f(x,y,z) = x^4 y + yz^3$, Let $x = rse^t$, $y = rs^2 e^{-t}$, $z = r^2 s(\sin(t))$

Repeat the exercise above:

i.e. Compute $\frac{\partial f}{\partial s}, \frac{\partial f}{\partial t}, \frac{\partial f}{\partial r}$ using chain rule and then w/o chain rule.

Recall from Calc I:

Given an equation involving both x, y:

(e.g. $(x-y)^2 = x + y^2$), we could compute "implicit derivatives".

$(x-y)^2 - x - y^2 = 0$

We said locally, $y = f(x)$, so we apply derivatives to obtain: $\frac{d}{dx}\left[(x - y(x))^2\right] = \frac{d}{dx}\left[x + (y(x))^2\right]$

Q: Why should that work?

$\overbrace{\text{Prop (Implicit Function Theorem)}}^{\text{IFT}}$: Let $F(x_1, x_2, \dots x_n)$ is diff and $\frac{\partial F}{\partial x_i}$ are cts on a disk about point P, and $\frac{\partial F}{\partial x_n}\Big|_{\vec{p}} \neq 0$, and $F(\vec{p}) = 0$, Then $x_n = f(x_1, x_2, \dots x_{n-1})$ is (near $\vec{p}$) a function of the other variables and $\frac{\partial f}{\partial x_i} = \left(-\frac{\partial F}{\partial x_i} \Big/ \frac{\partial F}{\partial x_n}\right)$